

**THE RELATIONSHIP OF GRADUATE RECORD EXAMINATION APTITUDE TEST
SCORES AND GRADUATE SCHOOL PERFORMANCE OF INTERNATIONAL
STUDENTS AT THE UNITED STATES UNIVERSITIES**

Ramazan Basturk

The Ohio State University

**A Paper Presented at the Annual Conference of the Mid-Western Educational Research
Association
Chicago, Illinois
October 13-16, 1999**

Purpose

One focus of educational research has been the examination of the relationship between students' entering academic characteristics and their subsequent achievement outcomes. In the case of admission tests, an important consideration is that test scores predict later achievement similarly for all groups of students who take the test (House et al., 1997). This paper will: (1) review the validity of the Graduate Record Examination for predicting international students' Graduate Grade Point Average (GGPA), and (2) inspect possible social, cultural and language bias or learning style differences in the prediction of international students' performance from Graduate Record Examination (GRE) test scores when used in admission / selection decisions at American universities.

Perspective and Theoretical Framework

Every year hundreds of international students apply to American universities for admission to graduate study. Only a small percentage of these students succeed. Indeed, the decision of whether to admit a student into U. S. graduate programs is an important challenge from both the students' and universities' perspective (Sternberg & Williams, 1997). For the students, the decision determines whether there is an opportunity for advanced education and establishes a good future career. For the universities, the admittance of poor-performing students may have an adverse effect on the school.

Currently, the large numbers of candidates seeking admission are forcing graduate admission committees to create a viable formula to accept or reject applicants into their graduate programs. Graduate admission committees primarily make their selection of prospective graduate students on the criteria of undergraduate grade point average (GPA) in the major field, recommendation letters from faculty members, Graduate Record Examination (GRE) or Graduate Management Admission Test (GMAT) tests scores depending on the graduate program, the written educational or career aspiration statement of the applicant, and previous work experience (Hamilton, 1990; Fisher & Resnick, 1990;

Abedi, 1991; Graham, 1991). However, the GMAT / GRE along with college grade point average (GPA) are commonly used as the first screen to eliminate all but the highest scoring individuals (Ingram, 1983; Goldenberg & Alliger, 1992; American Psychological Association, 1993; Educational Testing Service, 1995). This is due to the fact that most admission procedures involve the rank ordering of students based on GMAT or GRE scores and College GPA, with top-down selection. According to Goldenberg & Alliger (1992), individuals with high college GPAs, who receive only average or above average GRE scores, will have a low probability of being selected for admission. One of the difficult challenges for international students is that there are no explicit cutoff scores with the GRE or standards for entrance to help applicants recognize the need for compensating qualifications to gain admission in United States universities. What is or is not an acceptable GRE score varies from university to university. Desirable GRE scores for the Verbal plus Quantitative sections may be as high as 1200 (Huitema & Stein, 1993), but the mean for admission to doctoral programs is 1112 (Stroup & Benjamin, 1982).

Importance of Study

Every year hundreds of students from foreign countries enroll in American graduate schools and this trend will probably continue in the future. The decision to admit foreign students into a graduate program affects their advanced education opportunity and future career. The admission of foreign students to graduate school is a very complicated issue. Unlike their American counterparts, foreign students often lack proficiency in the English language due to differences in their native language, learning style and their cultural, social and economical background (Sharon, 1971, 1972; Craven, 1981; Kaiser, 1983; Wilson, 1986). The appraisal of the foreign candidate's aptitude for graduate study by standardized admission tests also has pitfalls. Poor performance may be due to factors not directly

related to aptitude for graduate study.

In order for foreign students in the US universities to be appropriately classified and their academic characteristics properly assessed, the admission committees need to create or be provided an effective evaluation system. In order to do that, confounding problems related to the GRE and its validity and usability with foreign students must be examined (Craven, 1981; Wilson, 1984a, 1986).

Validity Studies of the GRE Aptitude Test

Thornel and McCoy's (1985) study suggested a much stronger relationship between GRE and GGPA. Their findings revealed ($N = 582$ graduate students) an $r = .43$, $p < .05$.

Harvancik & Gordon (1986) collected data from 619 master's degree graduates to determine the extent of relationship between graduate grade point average and scores on the GRE. Their results were statistically significant. ($r = .48$, $p < .05$). House, Johnson and Tolone (1987) collected data from 76 graduate students in psychology and found $r = .43$, $p < .05$.

Other studies support the findings of a weak correlation between GRE and GGPA. Milner, King and McNeil (1984) drew their conclusion from a group of 145 full-time graduate students and found a low statistically significant relationship between GRE and GGPA ($r = .238$). Robertson and Nielsen (1961) also found a weak correlation ($r = .29$) between the GRE and faculty ratings of students' potential to complete a Ph.D. program. When using completion of a master's program as the criterion, House and Johnson (1993b) found no consistent pattern of predictability among the GRE scores, prior academic background, and graduate school performance. Ingram (1983) noted that "empirical evidence cannot support the validity of the GRE" (p. 711).

In 1992, Goldenberg and Alliger conducted a meta-analytic study of GRE and its validity in predicting success in graduate programs in psychology and counseling. The authors measured the

subscores of the GRE against graduate success, which include graduate grade point average (GGPA), comprehensive examination scores, whether graduated or not graduated, as well as specific course grades. They found that only the advanced psychology test of the GRE appeared to have positive correlation to graduate success. The verbal and quantitative scores of GRE only accounted for 3 % of the variance of GGPA.

It should be noted at this point that educational scholars have proposed different explanations for the low predictive validity results of the GRE aptitude test. First, Educational Testing Service (1992) argues that most validity studies on the GRE are biased because of sampling limitations. Holden (1993) supported and pointed out the same conclusion in his research. Since graduate programs often require minimum GRE scores for acceptance, subsequent analysis excludes persons with low academic potential, that is, people with substandard scores on the GRE. Not surprisingly, researchers who then gauge the performance of better students report that the GRE scores register low predictive validity; the homogeneity of the sample attenuates the predictive validity of the test. This reasoning suggests that if the samples were more heterogeneous the test would be more predictive (Oldfield & Hutchinson, 1997).

Second, range limitations in the output criterion can produce low validity coefficients. The graduate grade point average is generally used as a dependent variable in research on the validity of the GRE. Most graduate students receive an A or B in every class, leading to low dispersion within the output criterion. As with the sampling bias that results from excluding students with low GRE scores from the analysis, limiting the range for the dependent variable can also lower the test validity coefficients (Educational Testing Service, 1992; Oldfield & Hutchinson, 1997). Clearly, when both the predictor and criterion variables are homogeneous,

this situation can significantly reduce the predictive validity of the GRE.

Third, according to Oldfield (1994b), many intervening variables can influence validity studies for the GRE, including, for example, differences in the (a) faculty's grading standards and class assignments and (b) students' characteristics such as academic major, age, race and gender. Any combination of these and other factors can significantly affect the test's predictive validity. Schneider and Briel (cited in Oldfield & Hutchinson, 1997) call this the "noisy data" problem.

Finally, Oldfield and Hutchinson (1997) argue that the GRE is perhaps inherently inappropriate for gauging the traits necessary for success in graduate school. Oldfield (1995) stated that the GRE does not measure important personality attributes usually necessary for studying well in advanced education, such as maturity, persistence, communications and social skills.

International Students and GRE Aptitude Test

International students who constitute about 6% of the graduate student population in the United States were not included in these research studies. According to the Admission Offices in most American Universities, foreign students are also required to take the GMAT or the GRE aptitude test. Their scores are used in a variety of selection decisions that affect their career and future plans, such as graduate admissions, assistantship, fellowships or other types of scholarships (Sharon, 1971, 1972; Craven, 1981; Wilson, 1986). The students represent a variety of cultures, languages and educational backgrounds. As a group, they are more vulnerable to test bias than American minority groups who have many things in common with the majority group (Powers, 1980; Wilson, 1986).

Educational Testing Service (ETS) became aware of this possible test bias for the first time in 1963. They initiated a study of 637 foreign students. The results showed that foreign students scored

lower on the GRE than their American counterparts (Pitcher & Harvey, 1963). In this research, Verbal and Analytic scores of foreign students were 364 and 470 whereas American counterparts had 520 and 640, respectively. Educational Testing Service has studied the same subjects and found similar results each time (Sharon, 1971; Powers, 1980; Wilson, 1985, 1986). Here are some important research findings related to this subject.

Sharon (1972) compared foreign students' scores with their counterpart American students' scores on the GRE-Verbal and GRE-Quantitative sections (see Table 1). The mean GRE scores of foreign applicants indicate great discrepancies, relative to American students, between their verbal and quantitative abilities as measured by the GRE.

TABLE 1

Means and SD's for the Foreign and American Applicants on GRE

<u>Test</u>	<u>Foreign Applicants</u>		<u>American Applicants</u>	
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
GRE-Verbal	348	65	516	129
GRE-Quantitative	609	128	524	138

As a group, foreign students are more than one standard deviation below the mean on GRE-Verbal but more one-half of one standard deviation above the mean on GRE-Quantitative score. At this point, the author indicated that half of the foreign students are majoring in engineering, technology and mathematics which require extensive use of quantitative ability. The mean GRE-Quantitative score of those applicants was 670 as compared to the mean of all other majors.

The author conducted an initial analysis combining GRE-Verbal or GRE-Quantitative with TOEFL-Total in a linear multiple regression system. Since there was a possibility that different abilities would be required for success in different fields, the central prediction analysis were conducted by major field in those fields which had a sufficient number of subjects. Table 2 indicates the number and

percentage of subjects in each major field and their test scores and GPAs.

TABLE 2

Test and GPA Means by Major Field

<u>Major Filed</u>	<u>N</u>	<u>%</u>	<u>TOEFL</u>	<u>GRE</u>		<u>GPA</u>
				<u>V</u>	<u>Q</u>	
Eng.-Tech-Math	492	50	544	360	670	3.47
Natural Science	176	18	522	320	610	3.32
Social Science	307	32	534	343	511	3.31
Total	975	100	537	348	609	3.39

Table 3 indicates the average validities (weighted by the number of cases at each major department) of predictors and certain predictor composites by major field. It can be seen in Table 3 that the best single overall predictor is GRE-Quantitative with a validity coefficient of .32 for all subjects.

TABLE 3

Validity of Predictors and Predictor Composites by Major Field

<u>Major Field</u>	<u>GRE-V</u>	<u>GRE-Q</u>	<u>TOEFL</u>	<u>TOEFL & GRE-V</u>	<u>TOEFL & GRE-Q</u>
Eng-Tech-Math	.22	.39	.21	.23	.39
Natural Sciences	.41	.59	.39	.42	.61
Others	.35	.28	.39	.39	.39
All Subjects	.24	.32	.26	.27	.34

Sharon (1972) classified foreign students with their TOEFL scores as low, middle and high English proficiency groups and with their educational major field as “Engineering-Technology-Mathematics”, “Natural Science” and “Others”. This author found that the validities for some of subgroups are substantially higher than the corresponding validities for the total group. (Table 4)

In the major field of engineering, technology and mathematics the validity of GRE-Verbal is raised from .22 to .35 in the low proficiency group and .22 to .36 in the middle proficiency group. In the

same major field the validity of the GRE-Quantitative is raised from .39 to .56 in the low proficiency group. In the “Other” major field the validity of GRE-Verbal is raised from .35 to .44 in the middle proficiency group and that of GRE-Quantitative increased from .28 to .35 in the middle proficiency group and .28 to .37 in the high proficiency group. Thus, an English proficiency test such as TOEFL may raise the validity of the GRE aptitude tests in predicting foreign students’ graduate school grade point average. Perhaps the most important finding in this study is that, in general, foreign students appear to succeed in American graduate schools in spite of scoring more than one standard deviation below the mean of American students on the GRE-Verbal (Sharon, 1972).

TABLE 4

GRE Aptitude Test Validities for Low, Middle and High English Proficiency Groups

<u>Major Field</u>	<u>Low</u>		<u>Middle</u>		<u>High</u>	
	<u>GRE-V</u>	<u>GRE-Q</u>	<u>GRE-V</u>	<u>GRE-Q</u>	<u>GRE-V</u>	<u>GRE-Q</u>
Eng-Tech-Math	.35	.56	.36	.42	.21	.43
Others	.30	.25	.44	.35	.38	.37

Craven (1981), suggested in his research that cultural and linguistic differences should be taken into account when applying admissions criteria to foreign students seeking to enter American universities. Like Sharon’s (1971, 1972) and Wilson’s (1986) studies, Craven (1981) found low verbal aptitude score in the GRE and he hypothesized that the *speediness* of the GRE examination may be one of the many factors for poor performance of foreign students on the verbal section of the GRE. This author concluded that more research is needed on foreign students, particularly those studying in the humanities and social sciences whose aptitudes are more difficult to measure than those in mathematics and natural sciences (Craven, 1981).

Kaiser’s research (1983) on foreign students investigated the predictive validity of the GRE at

the University of Kansas. Different predictor variables such as GRE-Verbal and Quantitative scores, field of study, gender, year of initial enrollment were used to predict the criterion variable, the Graduate Grade Point Average (GGPA). The results from this study also confirmed that foreign students scored significantly lower than American students on the GRE scores, with poor correlation with the GGPA. This poor correlation between GRE scores and the criterion suggested to this author that the GRE is not the most appropriate way to predict the academic performance of foreign students. This author found that the high within-group variability among foreign students suggests the need to study subgroups formed on the basis of language, culture, or geographic location. In addition, the author pointed out that the use of separate norms for foreign students might be desirable (Kaiser, 1983).

Kenneth M. Wilson's report (1986), "The Relationship of GRE General Test Scores to First Year Grades for Foreign Students," is one of the basic related references about this topic. The author surveyed the validity of GRE scores for foreign students enrolled in the United States graduate schools, concentrating on three populations: (1) International Students who were not homogeneous with respect to linguistic, cultural and educational background; (2) Subgroups with homogeneous country of origin and background variables; and (3) Subgroups classified according to English proficiency, as indicated by Test of English as a Foreign Language (TOEFL) scores; GRE-Verbal, Analytic, and Quantitative scores; and self-reported English language proficiency. According to the author, for all subgroups of quantitatively-oriented departments, GRE-Quantitative scores and GRE-Analytical scores were more highly correlated with first year average grades (FYA) than were GRE-Verbal scores. Their correlation coefficients were .311, .275, and .097 for quantitative, analytical, and verbal scores, respectively (p. S-7). However, for the small sample of 85 students from five social science departments, verbal scores were most valid and quantitative scores were least valid. He found mean coefficient for verbal, analytical, and quantitative scores to be .253, .184, and .116, respectively (p. S-7).

The average scores of foreign students on the GRE-Verbal and Analytical ability measures (382 and 486, respectively) were more than one standard deviation lower than for U.S. examinees with the same quantitative ability (698 foreign students and 687 for U.S. students). In evaluating this finding, this author pointed out that foreign students generally recognize international mathematical symbol, and that the GRE-Quantitative section requires comparatively low levels of general English language verbal communication skill. The report concludes that the average quantitative performance of foreign students is absolutely comparable to that of U.S. examinees in similar fields of study because mathematics has its own language. On the other hand, their average performance on the GRE-Verbal and GRE-Analytic ability measures is significantly lower because of their less-than-native level of English proficiency (p. S-8). Even though measurement experts have always emphasized the need to make the language of the test a neutral vehicle, this may not be true for the GRE aptitude test when given to foreign students.

Undergraduate records, which generally have been found to be the best predictor of graduate success, are difficult when used with foreign students. (Sharon, 1971, 1972; Kaiser, 1983; Wilson, 1986). Their lack of comparability in the grading systems of universities in different countries makes it impossible to employ the prediction approach used with American students.

The appraisal of the foreign candidate's aptitude for graduate study by standardized admissions tests also has pitfalls. Poor performance may be due to factors not directly related to aptitude for graduate study. For example, non-native examinees may lack adequate English proficiency to understand the test questions or be unfamiliar with the philosophy or method of American objective tests (Sharon, 1972; Craven, 1981; Kaiser, 1983).

Competence in the English language is one factor that has been assumed to be crucial for the success of foreign students studying at an American university. It is very difficult to imagine how a student can learn in an American graduate school without being able to read, write, and comprehend in

the English language. Hence, English proficiency might be thought of as a necessary, even though not sufficient, prerequisite for graduate school success. (Sharon, 1972; Craven, 1981; Kaiser, 1983; Wilson, 1986). For this reason, many graduate schools recommend or require that their foreign applicants take the Michigan Test or Test of English as a Foreign Language (TOEFL) in their native country prior to coming or before starting the graduate program in the United States. Some graduate schools also require that their foreign applicants take the Test of Written English (TWE) and Test of Spoken English (TSE) to evaluate their writing ability and general oral language communication proficiency.

REFERENCES

- ABEDI, J. (1991) Predicting Graduate Success from Undergraduate Academic Performance: A Canonical Correlation Study. *Educational and Psychological Measurement*, 51, 151-160.
- AMERICAN PSYCHOLOGICAL ASSOCIATION. (1993). *Getting in: a step by step plan for gaining admission to graduate school in psychology*. Washington, DC: Author.
- CRAVEN, T. F. (1981) *A comparison of admission criteria and performance in graduate school for foreign and American students at Temple University*. Unpublished Ph.D. dissertation. Temple University, 1981.
- EDUCATIONAL TESTING SERVICE. (1992). *1992-93 guide to the use of the Graduate*

- Record Examination Program*. Princeton, NJ: Educational Testing Service.
- EDUCATIONAL TESTING SERVICE (1995). *GRE 1995-96 information and registration bulletin*. Princeton, NJ: Author.
- FISHER, J.B. & Resnick, D.A. (1990). Standardized Testing and Graduate Business School Admission: A review of Issues and an Analysis of a Baruch College MBA Cohort. *College and University*, 55, 137-148.
- GOLDENBERG, E. L., & Alliger, G.M. (1992) Assessing the Validity of the GRE for Students in Psychology: A Validity Generalization Approach. *Educational and Psychological Measurement*, 52, 1019-1027.
- GRAHAM, L.D. (1991) Predicting Academic Success of Students in a Master of Business Administration Program. *Educational and Psychological Measurement*, 51, 721-727.
- HAMILTON, D. J. (1990) Multiple Regression Analysis and Predicting of GPA upon Degree Completion. *College Student Journal*, 24, 91-96
- HARVANKIK, M.J. & Gordon, G. (1986) Graduate Record Examination Scores and Graduate Point Average: Is there a Relationship ? *American Association for Counseling and Development*, April, 21-23.
- HOLDEN, C. (1993) Predicting performance in graduate school. *Science*, 260, 494.
- HOUSE, D. J., Johnson, J.J. & Tolone, W.L. (1987) Predictive Validity of Graduate Record Examination Score for Performance in Selecting Graduate Psychology Courses. *Psychology Reports*, 60, 107-110.
- HOUSE, D. C., Johnson, J. J. (1993b). Predictive validity of the Graduate Record Examination advanced psychology test for graduate grades. *Psychological Reports*, 73, 184-186.
- HOUSE, J. D., Gupta, S & Xiao, B. (1997). Gender differences in prediction of grade

- performance from Graduate Record Examination. *ERIC Document no: ED 415 280*.
- HUTEMA, B. E., & STEIN, C., R. (1993). Validity of the GRE without restriction of range. *Psychological Reports, 72*, 123-127.
- INGRAM, R. E. (1983) The GRE in the Graduate Admission Process: Is How It Is Used Justified by the Evidence of Its Validity? *Professional Psychology: Research and Practice, 14*, 711-714.
- KAISER, J. (1983) The differential predictive validity of the GRE aptitude test for foreign students. *ERIC Document No: ED 227 174*.
- MILNER, M., McNeil, J.S. & King, S.W. (1984) The GRE: A question of Validity in Predicting Performance in Professional Schools of Social Work. *Educational and Psychological Measurement, 44*, 945-950.
- OLDFIELD, K. (1994a) On the importance of informing students about the potential risks associated with taking the Graduate Record Examination. *Journal of Thought, 29*, 61-70
- OLDFIELD, K. (1995) An impolite view of the Graduate Record Examination: why most studies find this test has low predictive validity. *Journal of Thought, 30*, 61-73
- OLDFIELD, K. & HUTCHINSON, J. R. (1997) Predictive validity of the Graduate Record Examination with and without range restraints'. *Psychological Reports, 81*, 211-220
- PITCHER, B., Harvey, P.R. (1963) The Relationship of GRE Aptitude Test Scores and Graduate School Performance of Foreign Students at Four American Graduate Schools. *GRE Report*, 63-1, Princeton, NJ.
- POWERS, D. E. (1980) The Relationship between Scores on the GMAT and the Test of English as a Foreign Language, *TOEFL Research Report, No:5*, Princeton, NJ.
- ROBERTSON, M., & NIELSEN, W. (1961) The Graduate Record Examination and the selection

- of graduate students. *American Psychologist*, 16, 648-650.
- SHARON, A. T. (1971) Test of English as a foreign language as a moderator of Graduate Record Examination scores in the prediction of foreign students' grades in graduate schools. *Research Report*, Educational Testing Service, NJ. ERIC Document No: 058 304
- SHARON, A. T. (1972) English Proficiency, Verbal Aptitude, and Foreign Students Success American Graduate Schools. *Educational and Psychological Measurement*, 32, 425-431
- STERNBERG, R. J. & WILLIAMS, W. M. (1997) Does the GRE predict meaningful success in the graduate training of psychologists? *American Psychologist*, June 1997, p. 630-641.
- STROUP, C. M., & BENJAMIN, L. T. (1982) Graduate study in psychology, 1970-79. *American Psychologist*, 37, 1186-1202.
- THORNEL, J.G., McCoy, A. (1985) the Predictive Validity of the GRE for Subgroups of Students in Different Academic Disciplines. *Educational and Psychological Measurement*, 45, 415-419.
- WILSON, K. M. (1984a) Foreign nationals taking the GRE general tests during 1981-82: selected characteristics and test performance. *GRE Professional Report No81-23bp & ETS RR-84-39*. Princeton, NJ: Educational Testing Service.
- WILSON, K.M. (1985) Factor Affecting GMAT Predictive Validity for Foreign MBA Students: An Exploratory Study. *Research Report*, ETS, Princeton, NJ.
- WILSON, K. M. (1986) The Relationship of GRE General Test Scores to First Year Grades for Foreign Graduate Students: *Educational testing Service, (ETS)*. Princeton, NJ.